The RunJob Project

Project Update

16-June-2004

Greg Graham, FNAL/CD

Outline of Talk

- Scope of project
- Technical Status
- Management Status
- Discussion

Scope of the RunJob Project

- Provides software to the experiments DZero, CDF, and CMS
  - Application configuration/workflow specification
  - Job building for multiple local environments or grids
  - Historically for large scale Monte Carlo production operations, but can be used for batch analysis operations also.
- Includes integration efforts "on project"
  - In return for an investment of manpower into the RunJob project, the project will manage integration tasks for the experiments.

Technical Status

- ShahKar 1.0 has been released
  - On schedule in mid May.
  - Includes all necessary functionality, and is being mainstreamed into CMS by end of June with a few weeks delay.
  - CMSShahkar.py integration module being developed for CMS could serve as a template for further integrations.

- ShahKar 1.1has been released in June
  - Includes up-to-date ScriptObjects subpackage which can be used by other software tools.
    - We had MOP and LCG (CMS) and D0Tools and SAMGrid (DZero) in mind.

Technical Status

- **CMS Integration Efforts**
  - Official CMS Production Tools "MCRunjob"
    - Includes "bread and butter" functionality: metadata specification of application configuration, contexts, scriptObjects, and canned production scripts.
  - DC04 Configurators designed for analysis of DST data in the recent CMS data challenge
    - prototypes, successfully used in Florida and FNAL
  - Monte Carlo Processing Service
    - being designed for USCMS as part of the DPE, will include Web Services decomposition MC request handling.

Technical Status

- DZero Integration
  - Integration going currently by Peter Love for D0 RunJob "Version 7"
    - but not mainstreamed!
  - ScriptObjects need to support a D0 RTE API.

- SAMGrid
  - ScriptObjects subpackage of ShahKar V1.1 ready
    - Ready for integration in D0Tools or "Version 6" with SAMGrid support
  - Use cases discussed at April 8 meeting.
    - Runtime configuration can be saved into ScriptObject automatically to provide a COMPLETE configuration record.

ScriptObjects

API
Basic Types:
-Key/Value Metadata
-Small Files or Scripts
-Designation of Unique Key
-Designation of Executable
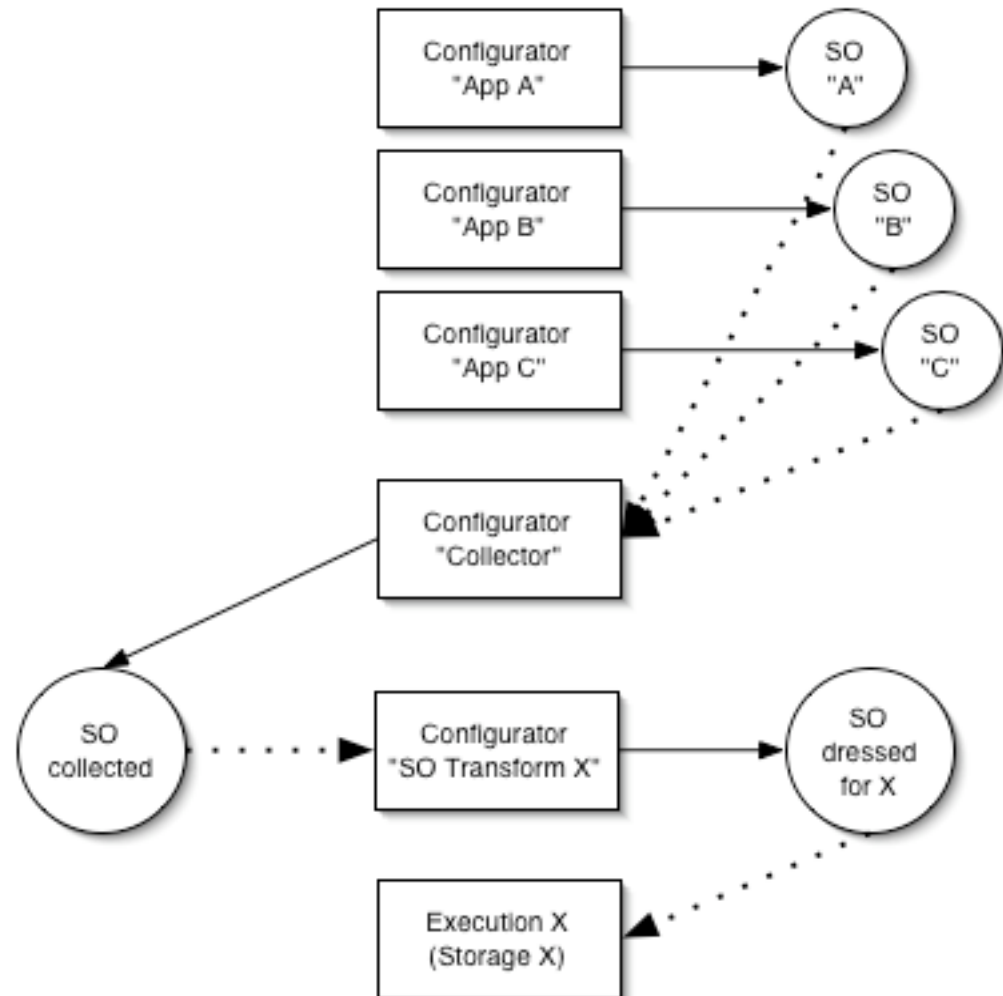**-Runtime Environment Variables**

Extension Types:
-File Brokers (API to native
facilities' file transport)
-Directory Brokers (Heirarchical
File Brokers)

Transforms on ScriptObjects:
-Additional files added
-Wrappers
-Archivers

16-June-2004   Greg Graham, FNAL/CD

Technical Status

- CDF Use of ShahKar
  - CAF Project: Needs work to ShahKar runtime extensions to handle conditional/cyclic workflows.
    - If we use SHREEK, we need to extend it.
    - At least, needs an API to another execution manager that does that if it exists.
  - Monte Carlo Production
    - We need to be able to process the TCL scripts that serve as application input.
- Both need an investment in manpower
  - and in management to identify manpower and coordinate with the project.

Technical Status

- Schedule:
  - We are roughly on track with core ShahKar planning as presented in February with two exceptions:
    - Integration effort on CDF, DZero
    - "Configurator Addressing"
      - We focussed on scriptObjetcs release instead!
  - Still, we need more feedback on the schedule.

16-June-2004   Greg Graham, FNAL/CD
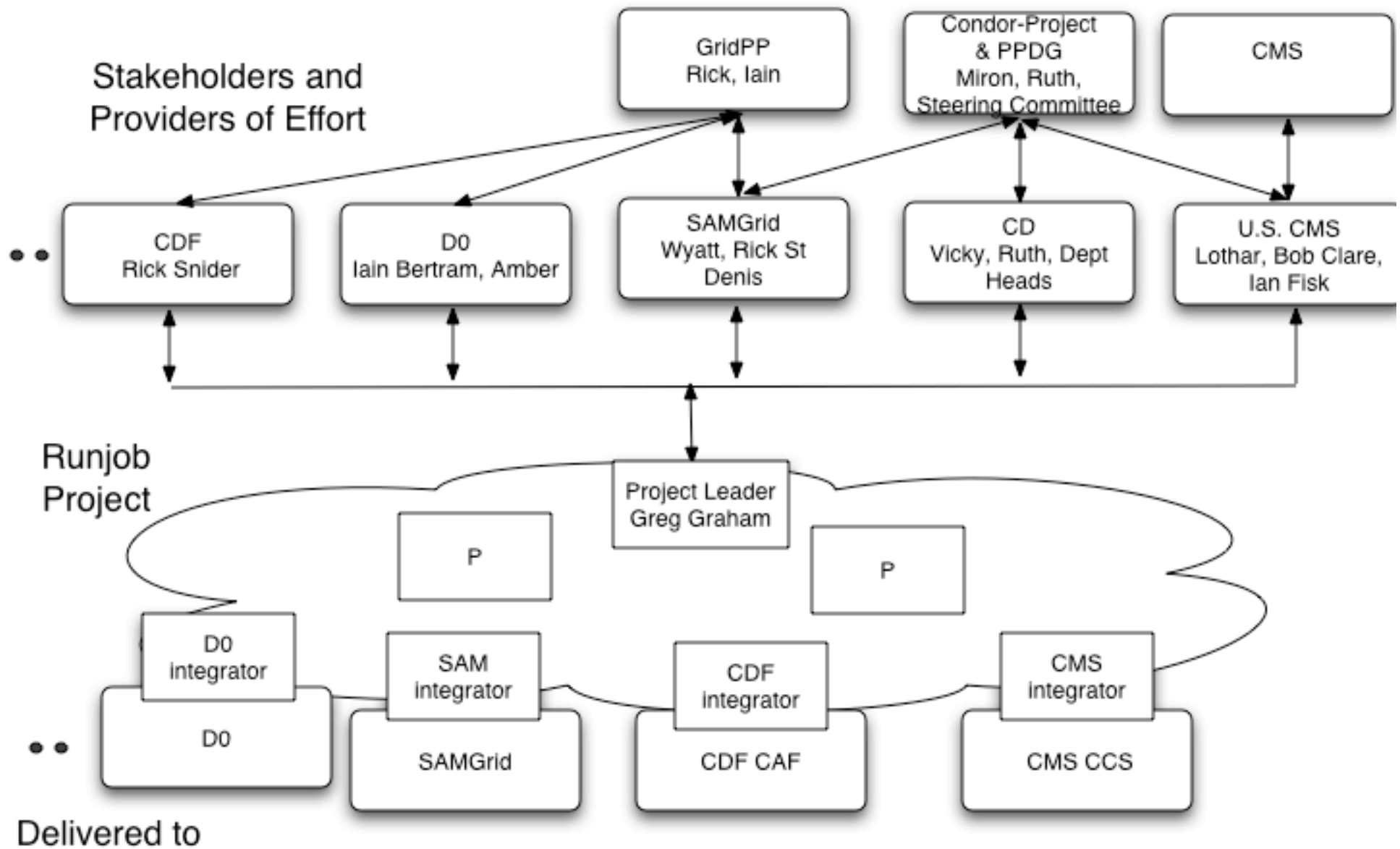
Management Status

- RunJob Project Staff
  - Eric Wicklund, 0.5 FTE
  - Greg Graham, Project Leader, 0.0 FTE
  - Dave Evans/G. Guglielmo, CMS Integration, 0.0 FTE
  - Peter Love, DZero Integration, 0.5 FTE (?)
- RunJob Effort Reporting is set up in the CD
- RunJob Project Page at http://projects.fnal.gov/runjob
- Technical Email list: shahkar@fnal.gov

Management Status

- Need a reliable structure to provide feedback from the experiments.
  - Management mailing list for stakeholders: project-runjob@fnal.gov
- I am willing to go to more experiment meetings
  - RunJob technical meetings are on every other Monday morning; but we clearly need more than these.
- "Status of project" should be sent out on a more regular basis, like before every GDM meeting.

16-June-2004   Greg Graham, FNAL/CD

Runjob Project Organization
5/5/04 draft

Stakeholders and Providers of Effort

| GridPP Rick, Iain | Condor-Project & PPDG Miron, Ruth, Steering Committee | CMS |

| CDF Rick Snider | D0 Iain Bertram, Amber | SAMGrid Wyatt, Rick St Denis | CD Vicky, Ruth, Dept Heads | U.S. CMS Lothar, Bob Clare, Ian Fisk |

Runjob Project

Project Leader Greg Graham

P

P

D0 integrator

SAM integrator

CDF integrator

CMS integrator

D0

SAMGrid

CDF CAF

CMS CCS

Delivered to

Discussion

Greg Graham, FNAL/CD